Fuzzified Choquet Integral With a Fuzzy-Valued Integrand and Its Application on Temperature Prediction

Rong Yang, Zhenyuan Wang, Pheng-Ann Heng, *Member, IEEE*, and Kwong-Sak Leung, *Senior Member, IEEE*

***Abstract*—In this paper, the original Choquet integral is generalized as a Fuzzified Choquet Integral with a Fuzzy-valued Integrand (FCIFI), which supports a fuzzy-valued integrand and an integration result. The calculation of the FCIFI is established on the Choquet integral with an interval-valued integrand (CIII). The definitions, properties, and calculation algorithms of the CIII and the FCIFI are discussed and proposed in this paper. As a specific application scheme, we designed a CIII regression model for the regression problems involving interval-valued data. This CIII regression model has a self-learning ability through a double genetic algorithm. Finally, a daily temperature predictor based on the CIII regression model is discussed, where a series of experiments is implemented to validate the performance of the predictor by real weather records from the Hong Kong Observatory.**

***Index Terms*—Choquet integral, fuzzy number, regression, temperature prediction.**

I. Introduction

As a classical nonlinear integral, a Choquet integral [4], [5], [14] with respect to a fuzzy measure or a signed fuzzy measure [11], [12] has been successfully performed as a nonlinear aggregation tool in information fusing and data mining for crisp databases [6], [10], [14], [18]. In these applications, the nonadditivity of the signed fuzzy measure provides an effective representation to describe the interaction among the contributions from the predictive attributes to the objective attribute.

The original Choquet integral only supports a real-valued integrand, which means that both the integrand and the integration result of the original Choquet integral are real valued. As such, the aforementioned data mining works based on the Choquet integral can only handle the problems concerning crisp real number. However, in many databases, some attributes may not be numerical, but categorical, or may have linguistic words (or may directly have fuzzy numbers) as their values [23]. Thus, to extend the advantages of the Choquet integral to the fuzzy domain such that it can manage fuzzy information, the original Choquet integral needs to be generalized (or fuzzified) such that it can be used to deal with fuzzy or linguistic data.

There are more than one way to fuzzify the Choquet integral. For a given signed fuzzy measure whose values are crisp real numbers, when the integrand is allowed to be fuzzy valued, we may define the integration result of its Choquet integral by either a crisp real number or a fuzzy number. The former is called the defuzzified Choquet integral with a fuzzy-valued integrand (DCIFI) [19], [21], which is named after its defuzzified (crisp real) integration result, whereas the latter is called the fuzzified Choquet integral with a fuzzy-valued integrand (FCIFI) [17], [20] due to its fuzzified (fuzzy-valued) integration result.

Both fuzzifications of the Choquet integral are applicable to different problems in the data mining area. The nonfuzzy integral result in the DCIFI facilitates to solve the classification [22] or clustering problems where crisp boundaries are pursued. On the other hand, the FCIFI is more suitable to the regression problems where the objective attribute is also fuzzy valued. In this paper, keeping the fuzzy knowledge in the integration result, we concentrate on the FCIFI.

In this paper, the theoretical derivation of the FCIFI is established based on the extension principle in the fuzzy set theory. Due to this definition, the calculation of the FCIFI has been transformed into that of the Choquet integral with an interval-valued integrand (CIII). A calculation scheme with a relevant algorithm is presented to obtain the value of the FCIFI with respect to a fuzzy measure and a signed fuzzy measure. Since one of the uncrisp data forms that frequently occur in databases is the interval data, a CIII regression model is designed for the regression problems where interval-valued data are involved. A double genetic algorithm (GA) optimization algorithm is derived for computing the internal regression coefficients of the CIII regression model, where one genetic approach is for the parameter optimization and the other approach is for the calculation of the CIII with respect to a signed fuzzy measure. The CIII regression model is utilized as a daily temperature predictor. In this predictor, the interval-valued weather indexes, such as the temperature range, relative humidity range, and gross rainfall, of several days before the day being considered act as the predictive attributes to forecast the temperature range of the day being considered by a CIII regression model. A series of experiments is implemented to validate the performance of the daily temperature predictor by real weather records from the Hong Kong Observatory.

Manuscript received May 28, 2007; revised August 24, 2007. This work was supported in part by the Natural Science Foundation of Guangdong Province, China, under Grant 06301289 and in part by the Shenzhen University Research and Development Fund under Project 200639. This paper was recommended by Associate Editor X. Wang.

R. Yang is with the Department of Automatic Science, College of Mechatronics and Control Engineering, Shenzhen University, Shenzhen 518060, China (e-mail: ryang@szu.edu.cn).

Z. Wang is with the Department of Mathematics, University of Nebraska, Omaha, NE 68182 USA (e-mail: zhenyuanwang@mail.unomaha.edu).

P.-A. Heng and K.-S. Leung are with the Department of Computer Science and Engineering, The Chinese University of Hong Kong, Shatin, Hong Kong (e-mail: pheng@cse.cuhk.edu.hk; ksleung@cse.cuhk.edu.hk).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSMCB.2007.911377

This paper is organized as follows. The basic concepts and properties of the original Choquet integral are reviewed in Section II. A formal definition of the FCIFI based on the extension principle is given in Section III. According to this definition, the calculation of the FCIFI is transformed into that of the CIII. In Section IV, the CIII acts as a new regression model for information fusing problems where not only crisp data but also interval data are involved. In Section V, the CIII regression model is used as a temperature predictor to forecast the temperature range of a day by a set of given weather records, such as temperature ranges, humidity ranges, cloud densities, or rainfalls, of the day before the day being considered. The experimental results and the comparisons with other existing approaches of the CIII daily temperature predictor are provided in Section VI. Finally, conclusions are given in Section VII.

II. Choquet Integral With a Real-Valued Integrand

Let $X = \{x_1, x_2, \ldots, x_n\}$ be a nonempty finite set of feature attributes and $\mathcal{P}(X)$ be the power set of X. A signed fuzzy measure [15] μ is a mapping from $\mathcal{P}(X)$ to $(-\infty, \infty)$ satisfying $\mu(\varnothing) = 0$. The set function μ is nonadditive in general, so it is also called a nonadditive set function. A fuzzy measure is considered to be a specialization of a signed fuzzy measure with additional monotonic property, i.e.,

$$\mu(A) \leq \mu(B) \qquad \forall A \subset B$$

on $\mathcal{P}(X)$. In this paper, for convenience, $\mu(\{x_{j_1}, x_{j_2}, \ldots, x_{j_m}\})$ is denoted by μ_j, where $j = \sum_{k=1}^{m} 2^{j_k - 1}$ and $\{j_1, j_2, \ldots, j_m\}$ is a subset of $\{1, 2, \ldots, n\}$. More explicitly, $\mu_0 = \mu(\varnothing) = 0$, $\mu_1 = \mu(\{x_1\})$, $\mu_2 = \mu(\{x_2\})$, $\mu_3 = \mu(\{x_1, x_2\})$, $\mu_4 = \mu(\{x_3\})$, $\mu_5 = \mu(\{x_1, x_3\})$, $\mu_6 = \mu(\{x_2, x_3\})$, $\mu_7 = \mu(\{x_1, x_2, x_3\}), \ldots$, and $\mu_{2^n - 1} = \mu(\{x_1, x_2, \ldots, x_n\})$.

Definition 2.1: Let $f : X \to (-\infty, \infty)$ be a real-valued function. The Choquet integral of f is defined as

$$\int f \, d\mu = \int_{-\infty}^{0} [\mu(F_\alpha) - \mu(X)] \, d\alpha + \int_{0}^{\infty} \mu(F_\alpha) d\alpha$$

where $F_\alpha = \{x | f(x) \geq \alpha\}$, for any $\alpha \in (-\infty, \infty)$, is the α-cut of f, which is represented as a crisp set of X.

For example, let $X = \{x_1, x_2, x_3\}$, and a real-valued function f is defined on X by $f(x_1) = 2.0$, $f(x_2) = 1.0$, and $f(x_3) = 3.0$; then, the α-cuts of f at $\alpha = 0.5, 1.5,$ and 2.5 are crisp sets of X, which are described by $F_{0.5} = \{x_1, x_2, x_3\}$, $F_{1.5} = \{x_1, x_3\}$, and $F_{2.5} = \{x_3\}$, respectively, as shown in Fig. 1.

To calculate the value of the Choquet integral of a given real-valued function f, usually, the values of f, i.e., $f(x_1), f(x_2), \ldots, f(x_n)$, should be sorted in nondecreasing order so that $f(x_1') \leq f(x_2') \leq \cdots \leq f(x_n')$, where $(x_1', x_2', \ldots, x_n')$ is a certain permutation of $\{x_1, x_2, \ldots, x_n\}$. Then, the value of the Choquet integral is obtained by

$$\int f \, d\mu = \sum_{i=1}^{n} \left[f(x_i') - f(x_{i-1}')\right] \cdot \mu\left(\{x_i', x_{i+1}', \ldots, x_n'\}\right) \quad (1)$$

where $f(x_0') = 0$.

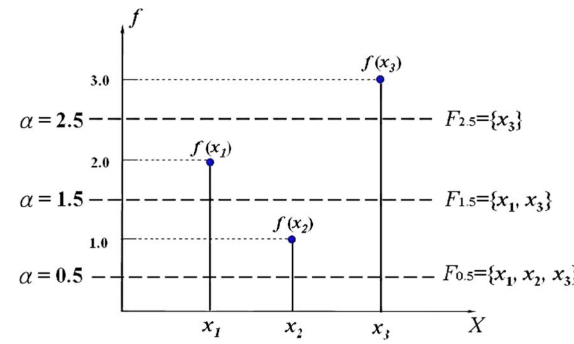


Fig. 1. α-cut of a real-valued function.

For convenience, Wang [14] has proposed a new algorithm to calculate the value of a Choquet integral with a real-valued integrand by the product of two vectors, as follows:

$$\int f \, d\mu = \sum_{j=1}^{2^n - 1} z_j \mu_j \quad (2)$$

in which

$$z_j = \begin{cases} \min\limits_{i: frc(\frac{j}{2^i}) \in [\frac{1}{2}, 1)} f(x_i) - \max\limits_{i: frc(\frac{j}{2^i}) \in [0, \frac{1}{2})} f(x_i), & \text{if it is} > 0 \text{ or } j = 2^n - 1 \\ 0, & \text{otherwise} \end{cases}$$
$$\text{for } j = 1, 2, \ldots, 2^n - 1$$

with a convention that the maximum on the empty set is zero, where $frc(j/2^i)$ is the fractional part of $(j/2^i)$. In the aforementioned formula, if we express j in the binary form $j_n, j_{n-1}, \ldots, j_1$, then $\{i | frc(j/2^i) \in [(1/2), 1)\} = \{i | j_i = 1\}$ and $\{i | frc(j/2^i) \in [0, (1/2))\} = \{i | j_i = 0\}$.

III. Fuzzified Choquet Integral With a Fuzzy-Valued Integrand

The Choquet integral with a real-valued integrand only supports a crisp-valued integrand; thus, it can only deal with crisp-valued data and is helpless when dealing with fuzzy information. To extend the advantages of the Choquet integral to fuzzy domain such that it can manage fuzzy information, fuzzifications of the original Choquet integral have been recently investigated [17], [21]. Such fuzzifications can support a fuzzy-valued integrand. They are regarded as generalizations of the original Choquet integral since they are able to handle diverse forms of information, including crisp data, interval values, fuzzy numbers, and linguistic variables. Fuzzifications of the Choquet integral can have their integration results fuzzified or defuzzified. In this paper, keeping the knowledge of the fuzziness in the integration result, we focus on the model with a fuzzy-valued integrand and a fuzzy-valued integration result, which is called the FCIFI.

First, we use the extension principle [2], [9] to define the Choquet integral with a measurable interval-valued integrand; that is, an integrand being a function whose range is a subset of $\mathcal{N}_R$ with the measurability in the following sense. Here, $\mathcal{N}_R$ denotes the set of all rectangular fuzzy numbers (interval numbers).

Definition 3.1: An interval-valued function $\bar{f} : X \to \mathcal{N}_R$ is measurable if both $f_l(x) = [\bar{f}(x)]_l$, which is the left endpoint of the interval $\bar{f}(x)$, and $f_r(x) = [\bar{f}(x)]_r$, which is the right endpoint of the interval $\bar{f}(x)$, are measurable functions of x.

Using a similar idea for the classical integrals in [1], we may define the Choquet integral of the interval-valued functions as follows.

Definition 3.2: Let $\bar{f} : X \to \mathcal{N}_R$ be a measurable interval-valued function on X and μ be a signed fuzzy measure on $\mathcal{P}(X)$. The Choquet integral of $\bar{f}$ with respect to μ is defined by

$$\int \bar{f}\, d\mu = \left\{ \int g\, d\mu \,|g(x) \in \bar{f}(x) \quad \forall x \in X\ , \right.$$
$$\left. g : X \to R \text{ is measurable} \right\}.$$

We call this Choquet integral as the CIII.

From Definition 3.2, we may directly obtain the following theorem.

Theorem 3.1: Let $\bar{f}$ and $\bar{g}$ be measurable interval-valued functions on $X = \{x_1, x_2, \ldots, x_n\}$; if $\bar{f}(x_i) \subseteq \bar{g}(x_i)$, $i = 1, 2, \ldots, n$, then $\int \bar{f}\, d\mu \subseteq \int \bar{g}\, d\mu$.

Using the idea of the representation theorem and extension principle [2], [9], we can now define the FCIFI. The integration result of this integral is a fuzzy subset of R.

Definition 3.3: A fuzzy-valued function $\tilde{f} : X \to \mathcal{N}$ is measurable if its α-cut, i.e., $\bar{f}_\alpha(x) = M_\alpha^{\tilde{f}(x)} = \{t | m_{\tilde{f}(x)}(t) \geq \alpha\}$, is a measurable interval-valued function for every $\alpha \in (0, 1]$, where $m_{\tilde{f}(x)}$ is the membership function of the value of $\tilde{f}$ at x. Here, $\mathcal{N}$ denotes the set of all fuzzy numbers.

Definition 3.4: Let $\tilde{f} : X \to \mathcal{N}$ be a measurable fuzzy-valued function on X and μ be a signed fuzzy measure on $\mathcal{P}(X)$. The Choquet integral of $\tilde{f}$ with respect to μ is defined by

$$\int \tilde{f}\, d\mu = \bigcup_{0 \leq \alpha \leq 1} \alpha \int \bar{f}_\alpha d\mu. \tag{3}$$

where $\bar{f}_\alpha(x)$ is given in Definition 3.3.

According to (3), the calculation of the Choquet integral with a fuzzy-valued function comes down to that of the Choquet integral with an interval-valued function. Here, we discuss this problem in two aspects, i.e., the Choquet integral of an interval-valued function with respect to a fuzzy measure and a signed fuzzy measure, respectively.

A. With Fuzzy Measure

Using the continuity and the monotonicity of the Choquet integral with the nonnegativity and the monotonicity of the fuzzy measures, we may prove the following theorem based on [7].

Theorem 3.2: Let $\bar{f} : X \to \mathcal{N}_R$ be a measurable interval-valued function on X and μ be a fuzzy measure on $\mathcal{P}(X)$. Then, the Choquet integral of $\bar{f}$ with respect to μ is

$$\int \bar{f}\, d\mu = \left[\int f_l d\mu, \int f_r d\mu \right] \tag{4}$$

where $f_l(x) = [\bar{f}(x)]_l$, i.e., the left endpoint of the interval $\bar{f}(x)$, and $f_r(x) = [\bar{f}(x)]_r$, i.e., the right endpoint of the interval $\bar{f}(x)$, $\forall x \in X$.

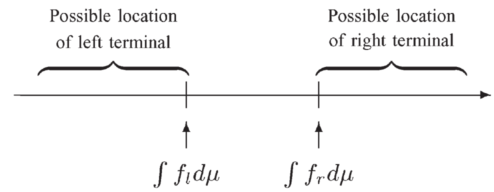


Fig. 2. Description of terminal ranges when μ is a signed fuzzy measure.

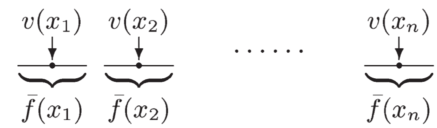


Fig. 3. Correspondence of function v and $\bar{f}(x_i)$.

B. With Signed Fuzzy Measure

We should note that the condition "Let μ be a fuzzy measure" cannot be replaced by the condition "Let μ be a signed fuzzy measure" in Theorem 3.2. This condition, which guarantees the monotonicity of μ, is essential. It can be verified by the following counterexample.

Example 3.1: Suppose that $X = \{x_1, x_2\}$, $\mu_1 = \mu(\{x_1\}) = 2$, $\mu_2 = \mu(\{x_2\}) = 3$, and $\mu_3 = \mu(X) = 1$. Then, μ is a signed fuzzy measure, not a fuzzy measure. Suppose there is an interval-valued function $\bar{f}$ that has values $[10, 12]$ at x_1 and $[8, 14]$ at x_2, we have $f_l(x_1) = 10$, $f_l(x_2) = 8$, $f_r(x_1) = 12$, and $f_r(x_2) = 14$. Then, we can obtain $\int f_l d\mu = 12$ and $\int f_r d\mu = 18$. Based on Theorem 3.2, we obtain $\int \bar{f}\, d\mu = [12, 18]$. However, the exact value of $\int \bar{f}\, d\mu$ is $\int \bar{f}\, d\mu = [10, 22]$.

As previously shown, with respect to a fuzzy measure μ, the left and right terminals of $\int \bar{f}\, d\mu$ can be directly calculated from the Choquet integrals of the integrand's left and right terminals, respectively. However, when the fuzzified Choquet integral is with respect to a signed fuzzy measure, Theorem 3.2 may not hold. In this case, the terminals of $\int \bar{f}\, d\mu$ may overstep the range, which is restricted by $\int f_l d\mu$ and $\int f_r d\mu$, as shown in Fig. 2. Hence, the exact membership function of the Choquet integral with respect to a signed fuzzy measure for a fuzzy-valued integrand is rather difficult to be found.

In this case, we need to find the actual left and right terminals of the CIII. This can be regarded as an optimization problem. To illustrate this optimization problem, we need to introduce a real-valued function $v : X \to (-\infty, \infty)$, where $v(x_i)$ is a number in the interval $\bar{f}(x_i)$, $i = 1, 2, \ldots, n$. The correspondence of function v and $\bar{f}(x_i)$ is depicted in Fig. 3.

We denote V as the set of all real-valued functions v. The optimization problem can now be summarized.

1) Finding a real-valued function $v_l : X \to (-\infty, \infty)$, where $v_l(x_i) \in \bar{f}(x_i)$, $i = 1, 2, \ldots, n$, so that

$$\int v_l d\mu = \min_{v \in V} \int v\, d\mu. \tag{5}$$

Here, $\int v_l d\mu$ is the left terminal of $\int \bar{f}\, d\mu$.

2) Similarly, finding another real-valued function $v_r : X \to (-\infty, \infty)$, where $v_r(x_i) \in \bar{f}(x_i)$, $i = 1, 2, \ldots, n$, so that

$$\int v_r d\mu = \max_{v \in V} \int v \, d\mu. \tag{6}$$

Here, $\int v_r d\mu$ is the right terminal of $\int \bar{f} \, d\mu$.

IV. Regression Model by CIII

Both the FCIFI and the CIII can be applied as regression tools. The former is a generalized model of the latter since the FCIFI handles heterogeneous fuzzy data whereas the CIII manages interval data, and as we know, interval data are included in heterogeneous fuzzy data. However, in this section, we focus on the latter because there are many practical cases where more complete information can be surely achieved by describing a set of variables in terms of interval data.

For example, intervals may occur as transaction time and valid time ranges in temporal databases, as line segments on a space-filling curve in spatial applications, as inaccurate measurements with tolerances in engineering databases, as daily temperatures registered as the minimum and maximum values, or for the minimum and maximum transaction prices daily recorded for a set of stocks.

The most widely used approaches on interval data analysis treat intervals as spread ranges with respect to a central value. The spread is generally assumed as the consequence of a measurement error and is considered as a perturbation in the data. However, the minimum and maximum of an interval value have their physical meanings and should not be simply regarded as measurement errors to the central value. We expect to handle them in a more reasonable manner in real applications.

A. CIII Regression Model

The classical Choquet integral has been proved to be a very powerful regression tool [10], [13], [16] because the nonadditivity of the fuzzy measure can thoroughly represent the nonlinear relationship between the predictive attributes and the objective attribute. Similarly, the CIII can also be regarded as a multiregression tool, which can represent the relationship not only among crisp data but also among interval data.

In our CIII regression model, let $x_1, x_2, \ldots, x_n$ be the predictive attributes and y be the objective attribute. Denote $X = \{x_1, x_2, \ldots, x_n\}$. The provided training data set consists of l observations of $x_1, x_2, \ldots, x_n$ and y and has the following form:

x_1	x_2	$\cdots$	x_n	y
$\bar{f}_{11}$	$\bar{f}_{12}$	$\cdots$	$\bar{f}_{1n}$	$\bar{y}_1$
$\bar{f}_{21}$	$\bar{f}_{22}$	$\cdots$	$\bar{f}_{2n}$	$\bar{y}_2$
$\vdots$				
$\bar{f}_{l1}$	$\bar{f}_{l2}$	$\cdots$	$\bar{f}_{ln}$	$\bar{y}_l$

where each row

$$\bar{f}_{j1} \quad \bar{f}_{j2} \quad \cdots \quad \bar{f}_{jn} \quad \bar{y}_j$$

is the jth observation of attributes $x_1, x_2, \ldots, x_n$ and y, $j = 1, 2, \ldots, l$. Note that the values of observations in the training data set are all interval numbers, denoted by bars on their heads. Positive integer l is the size of data and should be much larger than n. Usually, l is at least five times of 2^n. One observation of $x_1, x_2, \ldots, x_n$ can be regarded as an interval-valued function $\bar{f} : X \to \mathcal{N}_R$. Thus, the jth observation of $x_1, x_2, \ldots, x_n$ is denoted by $\bar{f}_j$, and we write $\bar{f}_{ji} = \bar{f}_j(x_i)$, $i = 1, 2, \ldots, n$, for $j = 1, 2, \ldots, l$. Similarly, the jth observation of y is denoted by $\bar{y}_j$, $j = 1, 2, \ldots, l$.

Hence, the CIII regression model is constructed as

$$\bar{y} = c + \int (a \cdot \bar{f} + b) d\mu$$

where

- $\bar{y}$ value of the objective attribute y;
- $\bar{f}$ interval-valued function on X with $\bar{f}(x_i)$ as its value at x_i, $i = 1, 2, \ldots, n$;
- μ signed fuzzy measure satisfying $\mu(\varnothing) = 0$;
- a real-valued function defined on X, which can be expressed as a scaling parameter vector $a = (a_1, a_2, \ldots, a_n)$;
- b real-valued function defined on X, which can be expressed as a shifting parameter vector $b = (b_1, b_2, \ldots, b_n)$;
- c adjusting constant.

The introduction of parameters $a_1, \ldots, a_n, b_1, \ldots, b_n$, and c attempts to balance the scales of the predictive attributes if in case they have different units. They are called by a joint name *regression coefficients* and should satisfy the following constraints:

$$-1 \le a_i \le 1 \text{ for } i = 1, 2, \ldots, n, \text{ with } \max_{1 \le i \le n} |a_i| = 1$$
$$b_i \ge 0 \text{ for } i = 1, 2, \ldots, n, \text{ with } \min_{1 \le i \le n} b_i = 0$$
$$c \in R.$$

In this multiregression model, all unknown parameters should be optimally determined before the regression model is put into operation. Here, we have $2^n + 2n$ unknown parameters to be learned: $2n + 1$ for the regression coefficients, and $2^n - 1$ for the values of the signed fuzzy measure. Our original scheme is to learn all these coefficients through a GA by describing them as genes in the chromosome. Nevertheless, as shown in Section III, when the CIII is with respect to a signed fuzzy measure, its integration result is calculated by a GA approach. Obviously, during the process of learning the coefficients of the CIII regression model, two GAs are involved. It will become a very time-consuming procedure if the number of the coefficients to be learned is not small. For this reason, simplification on the CIII regression model ought to be applied.

The new nonlinear interval-valued regression model is expressed as

$$\bar{y} = \int a \cdot \bar{f} \, d\mu$$

where only the scaling coefficients $a = (a_1, a_2, \ldots, a_n)$ are kept. Note that $(a \cdot \bar{f})$ is still an interval-valued function. The precondition of such simplification is that the training data set should be preprocessed. The interval values of the predictive attributes, as well as that of the objective attribute, are normalized into interval values in the range $[0, 1]$, respectively. The shifting coefficients $b = (b_1, b_2, \ldots, b_n)$ and the adjusting constant c are neglected because we assume that after normalization, the

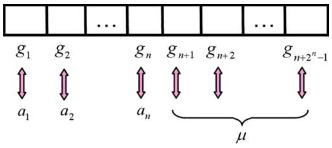


Fig. 4. Structure of an individual of chromosome in the double-GA optimization algorithm.

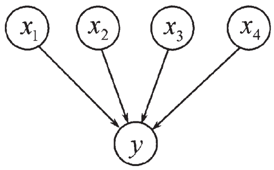


Fig. 5. Benchmark model in the explanatory experiments of the CIII regression model.

values of all the attributes are referred to the same origin, i.e., zero. Now, the number of the unknown parameters in the new CIII regression model is reduced to $n + 2^n - 1$.

B. Double-GA Optimization Algorithm

We propose a double-GA optimization algorithm to learn the unknown parameters in the CIII regression model. Remember that there are $n + 2^n - 1$ parameters to be determined: n for the scaling coefficients, and $2^n - 1$ for the values of μ. All of them are represented as genes of individuals of chromosome and decoded as real values between -1 and 1. Fig. 4 shows the structure of an individual represented in the double-GA optimization algorithm.

To evaluate the fitness value of an individual in the double GA, we define the distance between two interval numbers $\bar{s}$ and $\bar{t}$ as

$$|\bar{s} - \bar{t}| = \sqrt{(\bar{s}_l - \bar{t}_l)^2 + (\bar{s}_r - \bar{t}_r)^2} \tag{7}$$

where $\bar{s}_l$, $\bar{s}_r$, $\bar{t}_l$, and $\bar{t}_r$ are the left and right terminals of $\bar{s}$ and $\bar{t}$, respectively.

Then, the fitness value of an individual being considered in the population is defined as

$$\hat{\sigma}^2 = \frac{1}{l}\sum_{j=1}^{l} \left|\bar{y}_j - \bar{y}_j^*\right|^2 \tag{8}$$

where $\bar{y}_j^*$ is the calculated integration result of the CIII regression model, which is identified by the parameters represented by the current individual, with respect to the jth record of the predictive attributes, and $\bar{y}_j$ is the jth record of the objective attribute in the training data set.

We now present the procedure of the double GA.

1) Choose a large prime s as the seed for the random number generator. Set the value for each genetic parameter, as shown in the list that follows.

   p: The population size. It should be a large positive integer. Its default is 100.

   α and β: The probabilities used in a random switch to control the choice of genetic operators for producing offspring from the selected parents. They should satisfy the condition that $\alpha \geq 0$, $\beta \geq 0$, and $\alpha + \beta \leq 1$. Their defaults are 0.4 and 0.4, respectively.

   ε and δ: The small positive numbers used in the stopping controller. Their defaults are 10^{-6} and 10^{-10}, respectively.

   $maxIC$: The limit number of generations that have no successive significant improvement. Its default is 10.

   $maxGC$: The limit number of generations. Its default is 10 000.

2) Read the number of the predictive attributes n, the number of training samples l, and the training samples. Calculate

$$\hat{\sigma}_{\bar{y}}^2 = \frac{1}{l}\sum_{i=1}^{l} |\bar{y}_i -_{\text{avg}} \bar{y}_i|^2$$

   where

$$_{\text{avg}}\bar{y}_i = \frac{1}{l}\sum_{i=1}^{l} \bar{y}_i.$$

3) Randomly create an initial population that consists of p individuals of chromosome. Initialize the generation counter GC and the improvement counter IC by 0. Initialize $\hat{\sigma}_{\bar{y}}^2 \to m_0(\hat{\sigma}^2)$, where $m_0(\hat{\sigma}^2)$ stores the minimum fitness value of individuals in the closest previous generation.
4) Decode each individual in the population to obtain its corresponding scaling coefficients $a_1, a_2, \ldots, a_n$ and the values of the fuzzy measure $\mu_1, \mu_2, \ldots, \mu_{2^n-1}$.
5) For each individual in the current population, using the decoded regression coefficients $a_1, a_2, \ldots, a_n$ and $\mu_1, \mu_2, \ldots, \mu_{2^n-1}$, which cooperated with each record in the training data set, derive the calculated integration result of the CIII regression model represented by the current individual by

$$\bar{y}_j^* = \int a \cdot \bar{f}_j d\mu, \qquad j = 1, 2, \ldots, l.$$

TABLE I
GENETIC PARAMETERS IN THE EXPERIMENTS ON THE CIII REGRESSION MODEL WITH RESPECT TO A SIGNED FUZZY MEASURE

Parameters	Values	Remark
$PopSize_{integ}$	100	Population size in GA_{integ}
$MaxGen_{integ}$	50	Number of the maximum generation in GA_{integ}
IC_{integ}	20	The limit number of improvement counter in GA_{integ}
$PopSize_{reg}$	100	Population size in GA_{reg}
$MaxGen_{reg}$	10000	Number of the maximum generation in GA_{reg}
IC_{reg}	10	The limit number of improvement counter in GA_{reg}
$P1_{reg}$	5	The first dividing point in GA_{reg}
$P2_{reg}$	30	The second dividing point in GA_{reg}
P_{mut}	0.4	Probability to perform mutation operator
P_{cro}	0.4	Probability to perform crossover operator
ε	10e-6	Stopping controller for the whole program
δ	10e-10	Improvement marker for successive generations

TABLE II
RESULTS OF TEN TRIALS IN THE EXPERIMENTS ON THE CIII REGRESSION MODEL WITH RESPECT TO A FUZZY MEASURE

Data set	Set1	Set2	Set3	Set4	Set5
Minimum fitness value	0.000534724	4.51491e-05 Converge at generation 5752	0.00994598	4.89893e-05 Converge at generation 7104	0.00187487
Data set	Set6	Set7	Set8	Set9	Set10
Minimum fitness value	0.00119481	9.48671e-05	0.00031994	5.79011e-05 Converge at generation 9242	0.0148154

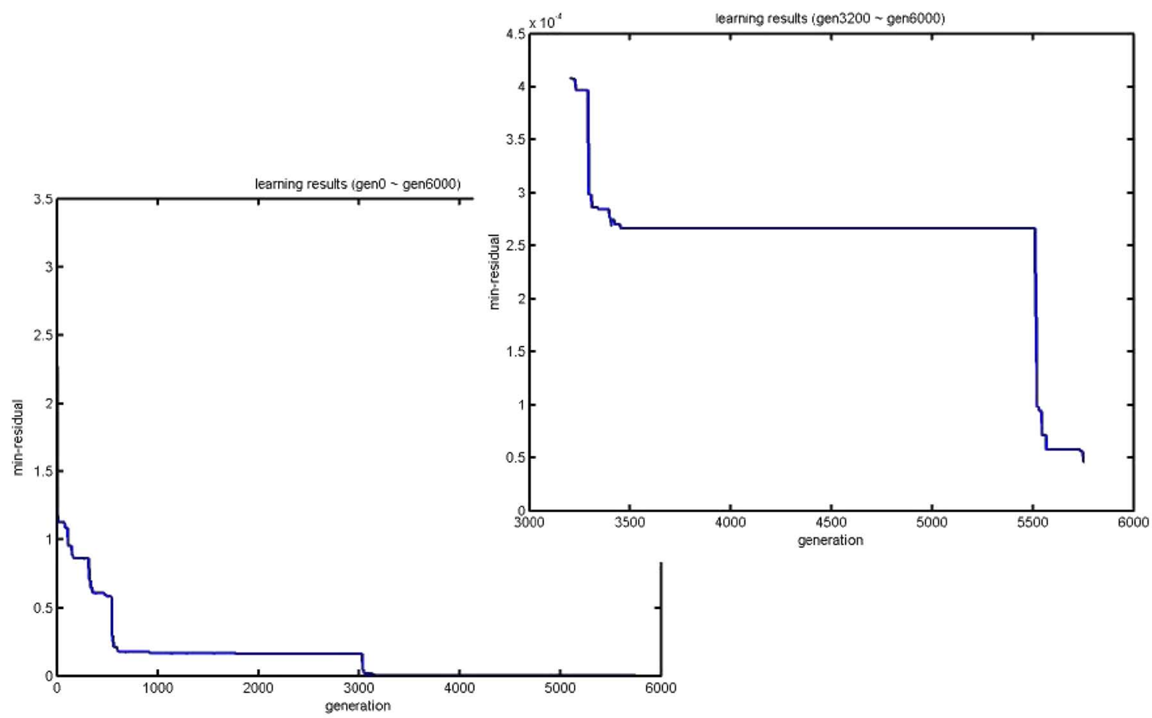


Fig. 6. Minimum fitness value versus iteration times in the best trial on the CIII regression model with respect to a signed fuzzy measure.

The approaches presented in Section III is applied to derive $\int a\bar{f}_j d\mu$. Then, the fitness value of the current individual is evaluated by (8).

6) The fitness value of the rth chromosome is denoted by $\hat{\sigma}_r^2$. Set $m(\hat{\sigma}^2) = \min_{1\le r\le p} \hat{\sigma}_r^2$, where $m(\hat{\sigma}^2)$ stores the minimum fitness value of individuals in the current generation.
7) If $m(\hat{\sigma}^2) < \varepsilon\hat{\sigma}_{\bar{y}}^2$ or $GC > maxGC$, then go to Step 13); otherwise, take the next step.
8) If $m_0(\hat{\sigma}^2) - m(\hat{\sigma}^2) < \delta\hat{\sigma}_{\bar{y}}^2$, then $IC + 1 \to IC$ and take the next step; otherwise, $0 \to IC$ and go to Step 10).
9) If $IC > maxIC$, divide the individuals in the current population into three parts by ascending order of their fitness values. The individuals in the first part are kept, whereas those in the second part create new offspring by random mutation, and those in the third part are replaced by new randomly created individuals of chromosome. Evaluate the new created individuals and update the population and then go to Step 12); otherwise, take the next step.
10) Do tournament selection (with a tournament size of 2). Randomly select one operator among the nonuniform mutation (with a probability of α), the BLX crossover (with a probability of β), and the random mutation (with a probability of $1 - \alpha - \beta$) to produce new individuals of chromosome as the offspring.
11) Repeat Step 10) until p new individuals are totally obtained. Evaluate this p new created individuals. Choose the best p individuals from the group of these p new created individuals and the original p individuals in the current generation to form the population for the next generation.
12) $GC + 1 \to GC$. Save $m(\hat{\sigma}^2)$ as $m_0(\hat{\sigma}^2)$. Then, go to Step 6).
13) Obtain the optimized $a_1, a_2, \ldots, a_n$ and $\mu_1, \mu_2, \ldots, \mu_{2^n-1}$ from the best individual in the current generation.
14) Stop.

C. Explanatory Experiments

Explanatory experiments have been implemented on synthetic data to verify the effectiveness and efficiency of the CIII regression model. Since the CIII regression model with respect to a signed fuzzy measure is a generalization of that with respect to a fuzzy measure, a series of experiments is conducted on the CIII regression model with a signed fuzzy measure. They all refer to a regression benchmark model with four predictive attributes and one objective attribute. Fig. 5 shows this benchmark model.

TABLE III
COMPARISONS OF THE PRESET AND THE ESTIMATED UNKNOWN PARAMETERS IN THE BEST TRIAL ON THE CIII REGRESSION MODEL WITH RESPECT TO A SIGNED FUZZY MEASURE

Coefficients	Preset value	Estimated value	Coefficients	Preset value	Estimated value
a_1	0.80	0.79122	$\mu(\{x_2, x_3\})$	0.50	0.57318
a_2	0.40	0.41223	$\mu(\{x_1, x_2, x_3\})$	0.60	0.59215
a_3	0.60	0.57989	$\mu(\{x_4\})$	0.20	0.19922
a_4	1.00	1.00000	$\mu(\{x_1, x_4\})$	-0.4	-0.40131
$\mu(\emptyset)$	0.00	0.00000	$\mu(\{x_2, x_4\})$	0.10	0.08827
$\mu(\{x_1\}$	0.10	0.10023	$\mu(\{x_1, x_2, x_4\})$	0.50	0.49993
$\mu(\{x_2\}$	0.20	0.19877	$\mu(\{x_3, x_4\})$	0.30	0.30177
$\mu(\{x_1, x_2\})$	-0.10	-0.10102	$\mu(\{x_1, x_3, x_4\})$	0.20	0.21004
$\mu(\{x_3\})$	0.40	0.41157	$\mu(\{x_2, x_3, x_4\})$	0.70	0.71049
$\mu(\{x_1, x_3\})$	0.30	0.29690	$\mu(\{x_1, x_2, x_3, x_4\})$	1.00	1.00000

By presetting the scaling coefficients a_1, a_2, a_3, and a_4 and the values of the signed fuzzy measure $\mu_1, \mu_2, \ldots, \mu_{15}$, ten training data sets, each of which consists of 200 observations, have been randomly generated for both experimental series, respectively.

For generalization, ten trials have been conducted to verify the performance of the proposed double-GA optimization algorithm on the CIII regression model with respect to a signed fuzzy measure. Since Theorem 3.2 does not work for this case, the genetic approach presented in Section IV-B is applied. The term "double GA" means one GA approach is dedicated to optimizing the unknown parameters of the CIII regression model, denoted by $\mathrm{GA}_{\mathrm{reg}}$, whereas another approach is devoted to calculating the integration result of the CIII with respect to a signed fuzzy measure, denoted by $\mathrm{GA}_{\mathrm{integ}}$. The relevant GA parameters that are involved are summarized in Table I.

We apply ten randomly generated data sets, each of which consists of 200 observations, to test the adaptability of our algorithm. The experimental results are recorded in Table II. Here, among the ten randomly generated training data sets, the trial on data set 2 gives the best optimization result. The optimization process stops at generation 5752 and converges to the optimal solution. For the remaining trials on other data sets, our double GA can also reach into the nearby space of the optimized point. This shows that our algorithm still has a satisfactory performance on the efficiency and effectiveness although double genetic approaches are involved.

Fig. 6 shows the convergence process of the trial based on data set 2. It reveals the change of the fitness value of the best individual in the population with respect to the iteration time. The comparisons of the preset and the estimated unknown parameters are listed in Table III. We can see that both the scaling coefficients and the values of the fuzzy measure have been well recovered.

V. CIII DAILY TEMPERATURE PREDICTOR

In this section, we discuss a practical application of the CIII regression model expressed in the previous sections. Here, we use the CIII regression model as a weather predictor to forecast the temperature range of a day through the meteorological information of several days before the day being considered.

Observing the daily meteorological data provided by the observatory, we find that interval values frequently occur. They may appear as the ranges of temperature, humidity bounds,

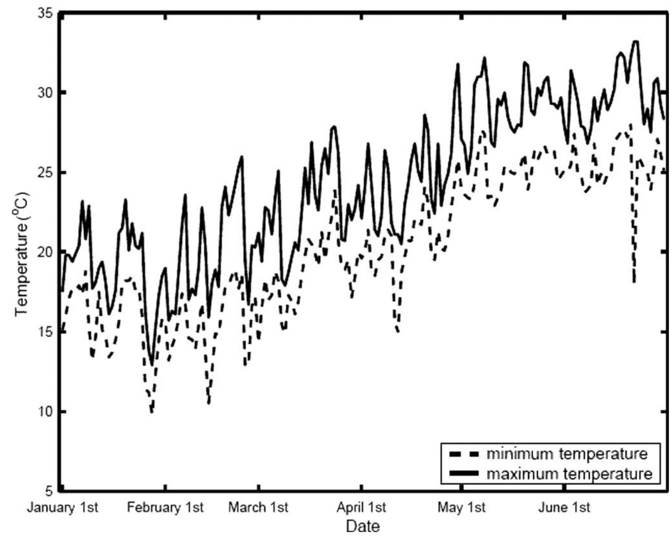


Fig. 7. Daily temperature ranges (in degrees Celsius) in Hong Kong from January to June 2006.

alterations of cloudage, or magnitudes of rainfall. Thus, directly managing the interval values is an essential requirement for the information fusing tools adopted for weather prediction. Furthermore, another important feature that those data mining tools should possess is the ability to elicit the internal relationship among the diverse weather indexes to meteorologists for deeper research and better model construction.

Here, we use the CIII regression model as a weather predictor to forecast the temperature range of a day through the meteorological information of several days before the day being considered. Among every day's weather records, we choose three indexes as the predictive factors that may influence the weather situation of the day being considered. They are the temperature range TR, which is represented as an interval value, the humidity range HR, which is also represented as an interval value, and the rainfall RF, which is represented as a crisp value. A window basis w containing the weather indexes of w days before the day being considered is used to predict the temperature range of the following day. For example, if $w = 2$, then six predictive factors are utilized to forecast the temperature range of the day being considered. They are TR_2, HR_2, and RF_2 of two days before the day being considered, and TR_1, HR_1, and RF_1 of one day before the day being considered. Obviously, the window basis w may range from 1 to 30 or even to a larger value. However, note that it is not necessarily true that a greater window basis brings a higher prediction accuracy. On the contrary, we will see in the following experiments that large values of w may somehow cause overtraining problems, which are similar to those in many data-driven models and ultimately influence the prediction accuracy. Moreover, it is an extremely time-consuming training process in the CIII daily temperature predictor if the window basis takes a large value.

A. Data Retrieval and Preprocessing

We collect the meteorological data for the whole year of 2006 in Hong Kong as the benchmark data of the CIII daily

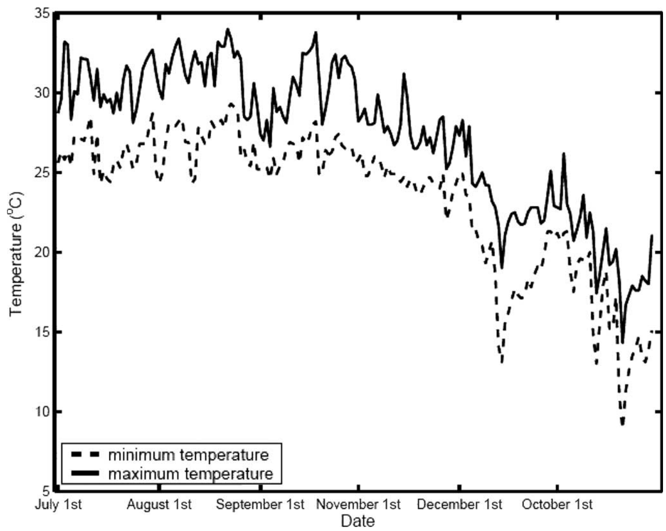


Fig. 8. Daily temperature ranges (in degrees Celsius) in Hong Kong from July to December 2006.

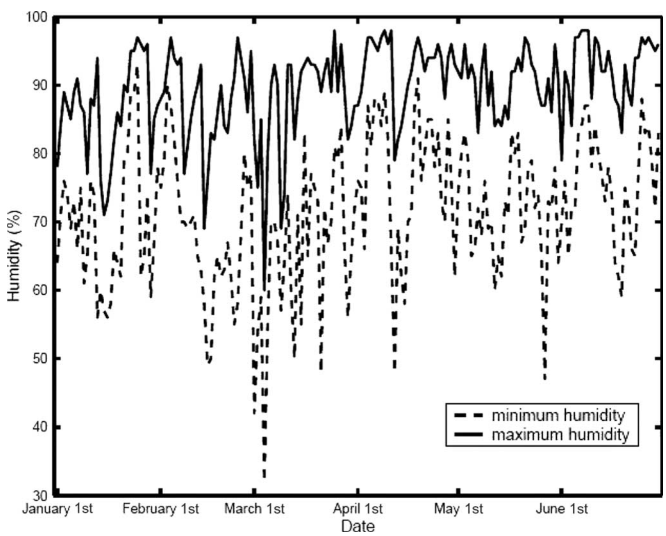


Fig. 9. Daily relative humidity ranges (in percentages) in Hong Kong from January to June 2006.

temperature predictor. Each observation of the data set is the record of some specific weather measurements of one day. These measurements include the following:

1) temperature range TR: an interval number with the left terminal as the minimum temperature and the right terminal as the maximum temperature of one day;
2) relative humidity range HR: an interval number with the left terminal as the minimum relative humidity and the right terminal as the maximum relative humidity of one day;
3) gross rainfall RF: a crisp number indicating the total rainfall of one day.

The daily temperature ranges (in degrees Celsius) in Hong Kong from January to June 2006 and July to December 2006 are plotted in Figs. 7 and 8, respectively. The daily relative humidity ranges (in percentages) in Hong Kong from January to June 2006 and July to December 2006 are plotted in Figs. 9 and 10, respectively. The daily gross rainfalls (in millimeters)

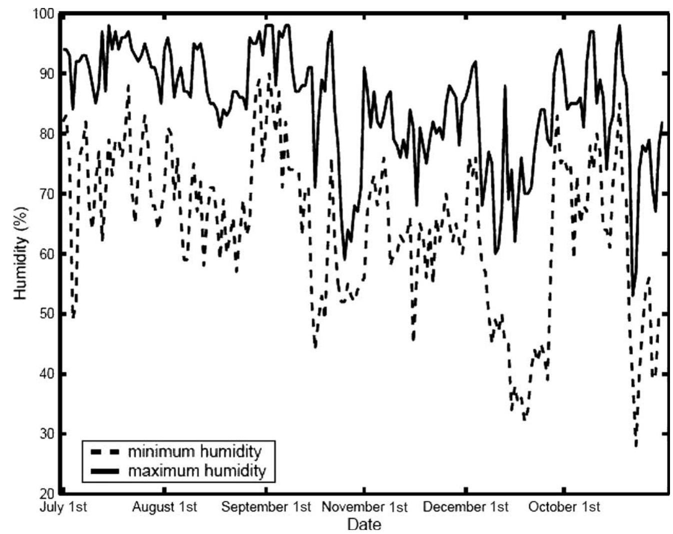


Fig. 10. Daily relative humidity ranges (in percentages) in Hong Kong from July to December 2006.

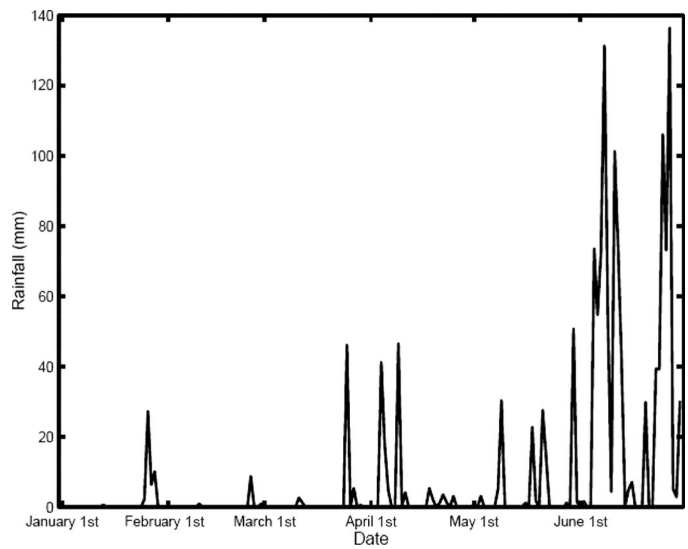


Fig. 11. Daily gross rainfalls (in millimeters) in Hong Kong from January to June 2006.

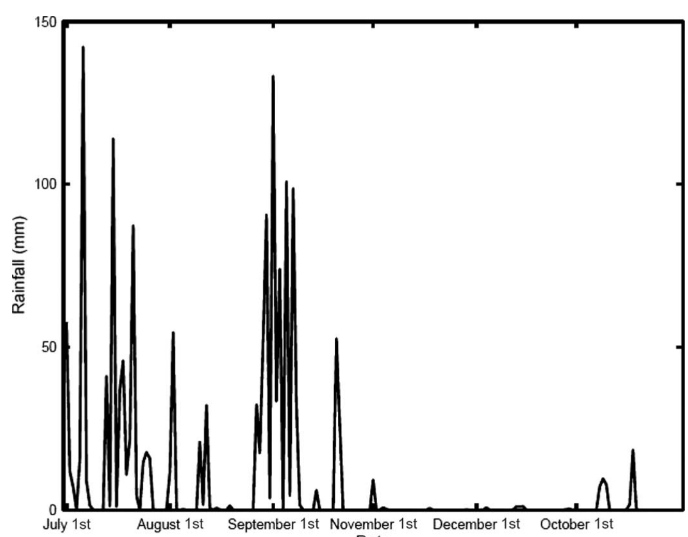


Fig. 12. Daily gross rainfalls (in millimeters) in Hong Kong from July to December 2006.

TABLE IV
DATA FORM OF THE TRAINING DATA IN THE CIII DAILY TEMPERATURE PREDICTOR

Predictive attributes										Objective attributes
Weather indices of 1 day before			Weather indices of 2 days before			...	Weather indices of w days before			Temperature range of current day
$\overline{TR}_{11}$	$\overline{HR}_{11}$	$\overline{RF}_{11}$	$\overline{TR}_{12}$	$\overline{HR}_{12}$	$\overline{RF}_{12}$	...	$\overline{TR}_{1w}$	$\overline{HR}_{1w}$	$\overline{RF}_{1w}$	$current\overline{TR}_1$
$\overline{TR}_{21}$	$\overline{HR}_{21}$	$\overline{RF}_{21}$	$\overline{TR}_{22}$	$\overline{HR}_{22}$	$\overline{RF}_{22}$	...	$\overline{TR}_{2w}$	$\overline{HR}_{2w}$	$\overline{RF}_{2w}$	$current\overline{TR}_2$
⋮										⋮
$\overline{TR}_{L1}$	$\overline{HR}_{L1}$	$\overline{RF}_{L1}$	$\overline{TR}_{L2}$	$\overline{HR}_{L2}$	$\overline{RF}_{L2}$	...	$\overline{TR}_{Lw}$	$\overline{HR}_{Lw}$	$\overline{RF}_{Lw}$	$current\overline{TR}_L$

in Hong Kong from January to June 2006 and July to December 2006 are plotted in Figs. 11 and 12, respectively.

To be managed by the CIII regression model, the raw weather data should be normalized into the range of $[0, 1]$. Take 50 and 0 as the upper and lower limits of the daily temperature in Hong Kong, respectively, the normalized daily temperature range can be obtained by

$$\overline{TR} = \frac{TR}{50.0} = \left[\frac{TR_l}{50.0} \quad \frac{TR_r}{50.0}\right] \tag{9}$$

where TR_l and TR_r denote the left and right terminals of the interval value TR, respectively. Similarly, considering 100 and 0 as the upper and lower limits of the daily relative humidity in Hong Kong, respectively, the normalized daily relative humidity range can be obtained by

$$\overline{HR} = \frac{HR}{100.0} = \left[\frac{HR_l}{100.0} \quad \frac{HR_r}{100.0}\right] \tag{10}$$

where HR_l and HR_r denote the left and right terminals of the interval value HR, respectively. Taking 200 and 0 as the upper and lower limits of the daily gross rainfall in Hong Kong, respectively, the normalized daily gross rainfall is

$$\overline{RF} = \frac{RF}{200.0}. \tag{11}$$

The normalized weather data are transformed into the training data sets for estimating the regression coefficients in the CIII daily temperature predictor. The training data set has the form shown in Table IV.

Here, the ith row records the temperature range, the relative humidity range, the gross rainfall of w days before the ith day, and the temperature range of the ith day, $i = 1, 2, \ldots, L$, where L is the number of records in the $\overline{TR}_{ij}$, $\overline{HR}_{ij}$, and $\overline{RF}_{ij}$ denote the temperature range, the relative humidity range, and the gross rainfall at the jth day before the ith day in the data set, respectively, and $current\overline{TR}_i$ denotes the temperature range of the ith day. Here, $i = 1, 2, \ldots, L$, and $j = 1, 2, \ldots, w$.

B. Variable Representation

Let $X = \{\overline{TR}_1, \overline{HR}_1, \overline{RF}_1, \overline{TR}_2, \overline{HR}_2, \overline{RF}_2, \ldots, \overline{TR}_w, \overline{HR}_w, \overline{RF}_w\}$ be a universal set, where w is a window basis that indicates that the relevant weather indexes of w days before the day being considered predict the temperature range of the day being considered. Let $\bar{f}$ be an interval-valued function mapping from X to an interval number or a crisp number (note that the crisp number is a special case of the interval number), and let $\mu : \mathcal{P}(X) \to (-\infty, \infty)$ be a signed fuzzy measure defined on X. The regression model of the CIII daily temperature predictor is defined by

$$current\overline{TR} = \int a\bar{f}\, d\mu \tag{12}$$

where a is expressed as a vector, i.e., $a = (a_{\overline{TR}_1}, a_{\overline{HR}_1}, a_{\overline{RF}_1}, a_{\overline{TR}_2}, a_{\overline{HR}_2}, a_{\overline{RF}_2}, \ldots, a_{\overline{TR}_w}, a_{\overline{HR}_w}, a_{\overline{RF}_w})$, to unify the different units of the weather indexes. Thus, the aforementioned model can predict the temperature range of the day being currently considered by the weather indexes of w days before it.

As stated in Section II, the fuzzy measure and the signed fuzzy measure in the Choquet integral can represent the strengths of the individual predictive attributes and any combination of them to determine the objective attribute. In the CIII daily temperature predictor, the values of the signed fuzzy measure also play an important role in representing the internal relationships among the weather indexes being considered. This

TABLE V
AVERAGE PREDICTION ACCURACY UNDER DIFFERENT VALUES OF THE WINDOW BASIS

Values of w	Average prediction accuracy (%)	
	Minimum temperature	Maximum temperature
1	87.433	86.421
2	98.791	98.673
3	94.204	94.968
4	90.369	91.753

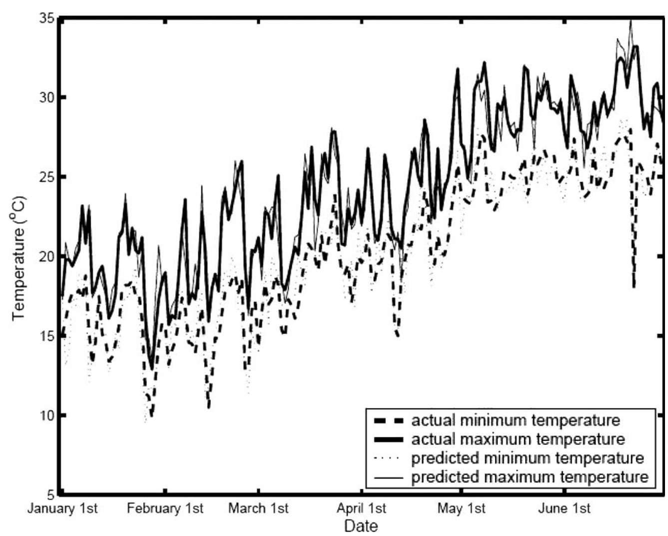


Fig. 13. Curve of the actual and predicted temperature ranges of the CIII daily temperature predictor with respect to a signed fuzzy measure when $w = 1$ (from January to June).

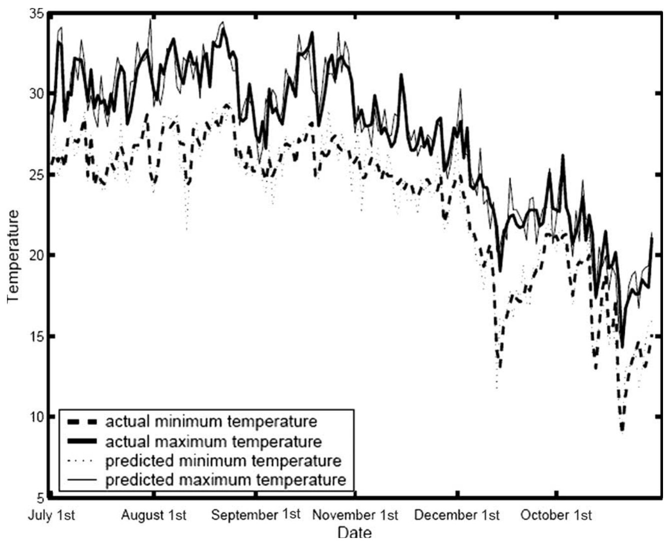


Fig. 14. Curve of the actual and predicted temperature ranges of the CIII daily temperature predictor with respect to a signed fuzzy measure when $w = 1$ (from July to December).

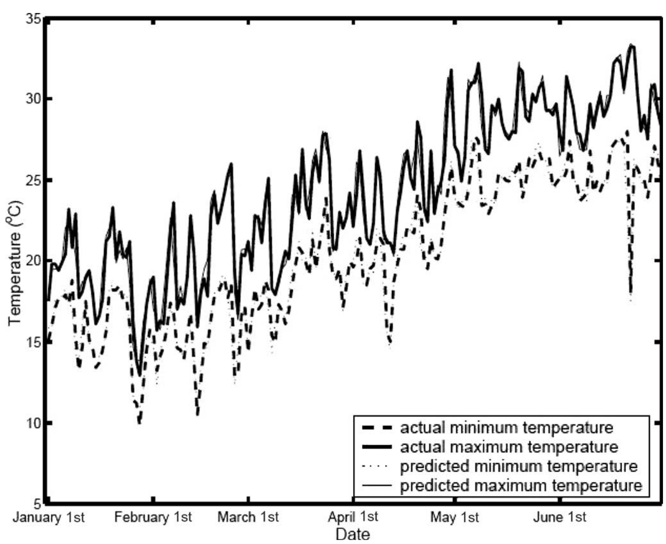


Fig. 15. Curve of the actual and predicted temperature ranges of the CIII daily temperature predictor with respect to a signed fuzzy measure when $w = 2$ (from January to June).

is of great benefit not only to improve the prediction accuracy but also to help meteorologists elicit the implicit internal correlations among different weather indexes. For more generalization, we choose a CIII regression model with respect to a signed fuzzy measure in the CIII daily temperature predictor. All regression coefficients (elements in vector a and values of μ) in the CIII daily temperature predictor are considered as unknown parameters and to be retrieved in a data-driven manner through the provided training data sets.

VI. Experiments and Comparisons

A. Experimental Results of the CIII Daily Temperature Predictor With Respect to a Signed Fuzzy Measure

In our implementations, the whole year's weather data (365 days) of Hong Kong in 2006 are used as the training data and the testing data. Experiments on different values of

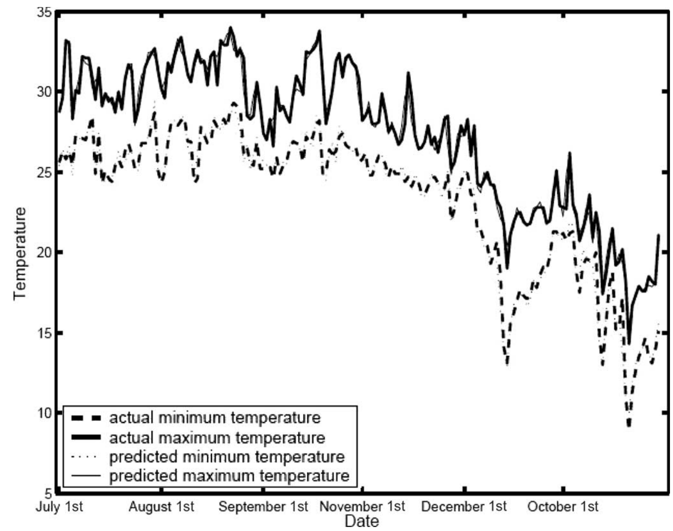


Fig. 16. Curve of the actual and predicted temperature ranges of the CIII daily temperature predictor with respect to a signed fuzzy measure when $w = 2$ (from July to December).

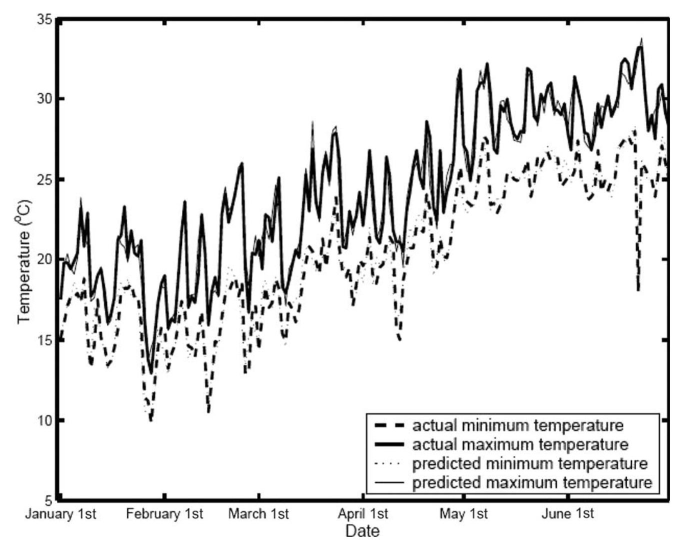


Fig. 17. Curve of the actual and predicted temperature ranges of the CIII daily temperature predictor with respect to a signed fuzzy measure when $w = 3$ (from January to June).

window basis, i.e., $w = 1$, $w = 2$, $w = 3$, and $w = 4$, have been conducted for ten times, respectively. The average prediction accuracies of the CIII daily temperature predictor under different window bases are listed in Table V, where the prediction accuracy is defined as

$$\Delta_l = \left(1 - \sqrt{\frac{1}{L}\sum_{i=1}^{L}\left|(current\overline{TR}_i)_l - (current\overline{TR}_i^*)_l\right|^2}\right) \times 100\%$$

for the minimum temperature and

$$\Delta_r = \left(1 - \sqrt{\frac{1}{L}\sum_{i=1}^{L}\left|(current\overline{TR}_i)_r - (current\overline{TR}_i^*)_r\right|^2}\right) \times 100\%$$

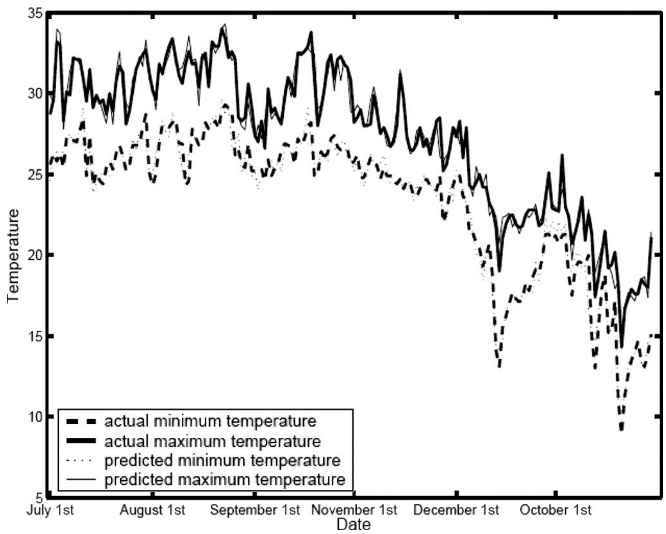


Fig. 18. Curve of the actual and predicted temperature ranges of the CIII daily temperature predictor with respect to a signed fuzzy measure when $w = 3$ (from July to December).

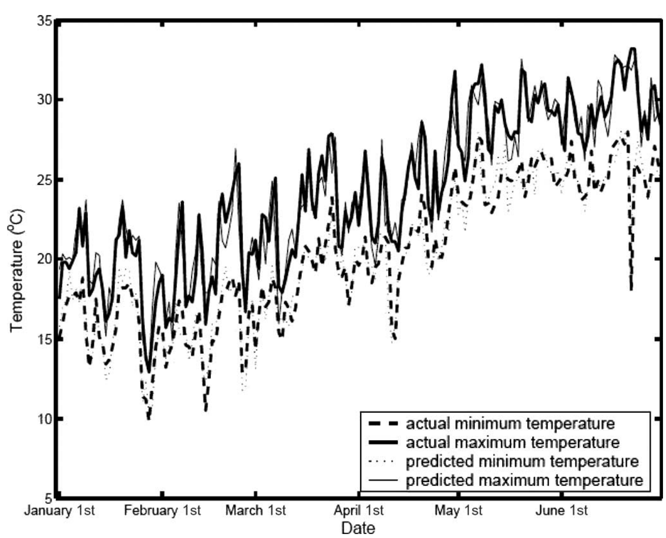


Fig. 19. Curve of the actual and predicted temperature ranges of the CIII daily temperature predictor with respect to a signed fuzzy measure when $w = 4$ (from January to June).

for the maximum temperature, where L indicates the number of records in the data set.

Here, $(current\overline{TR}_i)_l^*$ and $(current\overline{TR}_i)_r^*$ denote the minimum and maximum values of the calculated temperature range of the CIII regression model at the ith day. $(current\overline{TR}_i)_l$ and $(current\overline{TR}_i)_r$ denote the minimum and maximum values of the actual temperature range of the ith day, respectively. Figs. 13–20 show the predictive and actual temperature ranges under different values of the window basis.

Among the series of experiments on different values of the window basis, the trial when $w = 2$ retrieves the best average prediction accuracy. It can predict the minimum and maximum temperatures by accuracies up to 98.791% and 98.673%, respectively, which are significantly better than the trials with other values of the window basis. Table VI lists the estimated regression coefficients (elements in vector a and values of μ) in the trained CIII daily temperature predictor

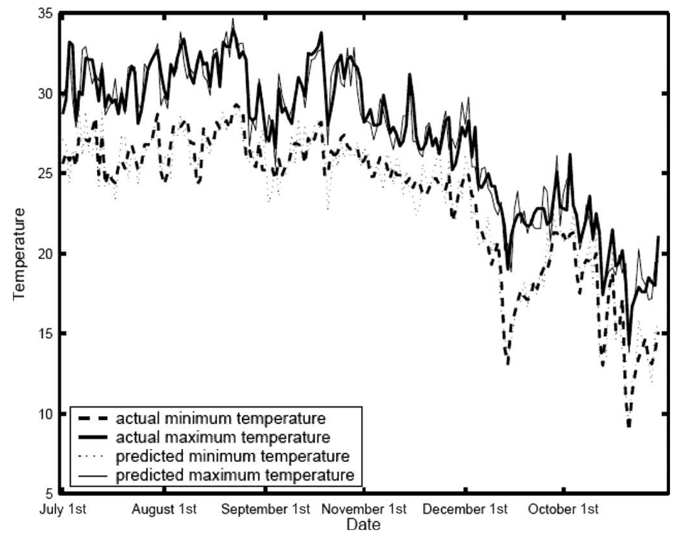


Fig. 20. Curve of the actual and predicted temperature ranges of the CIII daily temperature predictor with respect to a signed fuzzy measure when $w = 4$ (from July to December).

TABLE VI
REGRESSION COEFFICIENTS (ELEMENTS IN VECTOR a AND VALUES OF μ) IN THE TRAINED CIII DAILY TEMPERATURE PREDICTOR WHEN $w = 2$

Coefficients	Values	Coefficients	Values
$a_{\overline{TR}_1}$	1.0	μ_{29}	0.2514
$a_{\overline{HR}_1}$	0.4822	μ_{30}	0.1547
$a_{\overline{RF}_1}$	0.2541	μ_{31}	0.2144
$a_{\overline{TR}_2}$	0.4925	μ_{32}	0.3523
$a_{\overline{HR}_2}$	0.4925	μ_{33}	0.0948
$a_{\overline{RF}_2}$	0.2497	μ_{34}	0.1672
μ_0	0.0	μ_{35}	0.4268
μ_1	0.8501	μ_{36}	0.1636
μ_2	0.6068	μ_{37}	0.2812
μ_3	0.2311	μ_{38}	0.4873
μ_4	0.4860	μ_{39}	0.7064
μ_5	0.7621	μ_{40}	0.3629
μ_6	0.4565	μ_{41}	0.5058
μ_7	0.8214	μ_{42}	0.2318
μ_8	0.7919	μ_{43}	0.1701
μ_9	0.1763	μ_{44}	0.1029
μ_{10}	0.2051	μ_{45}	0.2838
μ_{11}	0.3468	μ_{46}	0.6472
μ_{12}	0.0289	μ_{47}	0.8522
μ_{13}	0.1764	μ_{48}	0.1472
μ_{14}	0.2066	μ_{49}	0.2573
μ_{15}	0.4734	μ_{50}	0.3118
μ_{16}	0.5225	μ_{51}	0.2541
μ_{17}	0.0176	μ_{52}	0.4290
μ_{18}	0.1361	μ_{53}	0.3270
μ_{19}	0.0994	μ_{54}	0.2031
μ_{20}	0.2330	μ_{55}	0.1795
μ_{21}	0.2093	μ_{56}	0.2856
μ_{22}	0.1231	μ_{57}	0.1008
μ_{23}	0.2626	μ_{58}	0.3450
μ_{24}	0.1013	μ_{59}	0.4078
μ_{25}	0.3361	μ_{60}	0.3560
μ_{26}	0.3191	μ_{61}	0.6451
μ_{27}	0.2098	μ_{62}	0.6343
μ_{28}	0.1897	μ_{63}	0.9216

when $w = 2$. Here, the estimated values of μ represent the individual and joint prediction efficiencies of the weather indexes $(\overline{TR}_1, \overline{HR}_1, \overline{RF}_1, \overline{TR}_2, \overline{HR}_2, \overline{RF}_2)$ to the temperature range of the following day. For example, $\mu_1 = \mu(\{\overline{TR}_1\}) = 0.8501$, $\mu_2 = \mu(\{\overline{HR}_1\}) = 0.6068$, $\mu_4 = \mu(\{\overline{RF}_1\}) = 0.4860$, $\mu_8 = \mu(\{\overline{TR}_2\}) = 0.7919$, $\mu_{16} = \mu(\{\overline{HR}_2\}) = 0.5225$, and $\mu_{32} = \mu(\{\overline{RF}_2\}) = 0.3523$ indicate the individual prediction efficiencies of six weather indexes, respectively. A small value at $\mu_{33} = 0.0948$ indicates

TABLE VII
AVERAGE PREDICTION ACCURACIES FOR THE MINIMUM AND MAXIMUM TEMPERATURES OF JANUARY 2007 AND FEBRUARY 2007 IN HONG KONG FROM DIFFERENT MODELS

Model	Average prediction accuracy (%)							
	Jan. 2007		Feb. 2007		Mar. 2007		Apr. 2007	
	Minimum temperature	Maximum temperature	Minimum temperature	Maximum temperature	Minimum temperature	Maximum temperature	Minimum temperature	Maximum temperature
Back Propagation	94.2689	93.8932	94.0140	93.6087	93.7379	94.0435	94.4815	93.6786
SND-gray-regression-fuzzy-model	95.3846	95.2081	94.8142	95.4885	95.2132	94.5774	94.5997	94.6238
Fuzzy time series	96.3564	95.2446	97.1668	96.4884	95.4297	97.2773	96.9806	96.8336
CIII temperature predictor	98.3013	97.8962	98.2278	98.3012	98.4847	97.7532	98.0549	97.5682

TABLE VIII
HISTORICAL DATA OF THE DAILY WEATHER INDEXES OF JANUARY 2007 IN HONG KONG

Day	Temperature range		Humidity range		Rainfall
	Minimum	Maximum	Minimum	Maximum	
1	17.4	19.6	65.0	84.0	0
2	17.8	23.4	52.0	83.0	0
3	17.1	19.6	61.0	83.0	0
4	17.8	20.5	61.0	83.0	0.01
5	18.3	21.8	62.0	89.0	0
6	18.6	20.9	76.0	88.0	0
7	18.1	20.2	71.0	88.0	0
8	17.5	19.7	65.0	84.0	0
9	17.2	20.0	62.0	84.0	0
10	18.2	19.5	74.0	86.0	0.01
11	17.0	20.9	65.0	95.0	1.8
12	14.2	18.5	56.0	73.0	0
13	11.3	16.5	58.0	71.0	0
14	14.5	16.9	58.0	80.0	0
15	16.1	19.6	74.0	85.0	0
16	16.1	22.3	74.0	91.0	0.3
17	13.7	17.0	65.0	78.0	0
18	13.6	16.8	75.0	95.0	15
19	10.5	13.7	86.0	97.0	30.7
20	8.8	11.9	82.0	94.0	2.5
21	8.0	12.9	43.0	77.0	0
22	9.1	13.7	59.0	71.0	0
23	9.0	13.8	59.0	84.0	0.7
24	9.0	12.7	34.0	62.0	0.01
25	8.5	13.1	43.0	65.0	0
26	12.0	14.5	51.0	73.0	0
27	12.7	14.7	56.0	76.0	0
28	12.3	14.7	58.0	84.0	0
29	14.2	16.6	76.0	86.0	0
30	15.3	19.7	74.0	89.0	0.01
31	15.5	17.6	77.0	90.0	0.01

TABLE IX
HISTORICAL DATA OF THE DAILY WEATHER INDEXES OF FEBRUARY 2007 IN HONG KONG

Day	Temperature range		Humidity range		Rainfall
	Minimum	Maximum	Minimum	Maximum	
1	16.2	19.4	82.0	90.0	0
2	18.0	21.8	78.0	92.0	0.01
3	11.0	18.0	76.0	94.0	1.7
4	8.6	11.2	76.0	96.0	15
5	7.8	11.7	79.0	97.0	14.9
6	11.4	12.8	64.0	82.0	0.01
7	8.9	12.7	71.0	97.0	11.5
8	8.3	10.3	87.0	97.0	8.7
9	7.7	14.5	66.0	85.0	0
10	11.9	16.3	51.0	89.0	0
11	13.2	17.8	56.0	80.0	0
12	15.4	21.2	60.0	83.0	0
13	15.9	21.3	41.0	84.0	0
14	15.8	21.6	45.0	79.0	0
15	16.4	20.4	42.0	83.0	0
16	17.2	20.2	63.0	81.0	0.01
17	17.9	23.1	63.0	85.0	0
18	17.9	20.5	78.0	91.0	0
19	17.7	22.8	72.0	91.0	0
20	19.2	22.9	74.0	90.0	0
21	19.2	22.9	74.0	90.0	0
22	19.4	24.8	65.0	92.0	0
23	17.9	21.7	62.0	78.0	0
24	17.1	19.7	72.0	83.0	0
25	19.2	23.0	68.0	83.0	0
26	19.1	21.9	72.0	88.0	0.01
27	18.2	21.7	66.0	85.0	0
28	19.3	22.8	77.0	88.0	0
29	20.6	24.7	72.0	93.0	0

that the weather indexes $\overline{TR}_1$ and $\overline{RF}_2$ have a low joint prediction efficiency in temperature forecasting. On the contrary, a large value at $\mu_{47} = 0.8522$ shows that the weather indexes $\overline{TR}_1$, $\overline{HR}_1$, $\overline{RF}_1$, $\overline{TR}_2$, and $\overline{RF}_2$ have a high joint prediction efficiency in temperature forecasting.

B. Comparisons and Discussions

To evaluate the performance of the CIII daily temperature predictor, we compare its forecasting results at $w = 2$ with those of other approaches, such as the hybrid gray-based model [8], the fuzzy time series model [3], and the backpropagation model of the neural network. Since these comparison approaches cannot directly deal with the interval data, the interval-valued weather indexes such as the temperature range and the relative humidity range are decomposed into two predictive attributes, respectively, such as the minimum temperature and the maximum temperature and the minimum relative humidity and the maximum relative humidity. Similarly, the temperature range to be predicted is also decomposed into two outputs, i.e., the minimum predictive temperature and the maximum predictive temperature. Note that in the following comparison experiments, our CIII temperature daily predictor uses the interval-valued data (both the minimum and maximum values) as the information source, whereas the comparison approaches use the crisp-valued data (either the minimum or maximum value). These kinds of comparisons appear to be a bit unfair since our approach has more information than those being compared. However, relative comparisons are still reasonable to be implemented for evaluating the performance of our predictor.

The training data set is still the collection of the whole year's weather data of Hong Kong in 2006, whereas the testing data are the collection of the historical data of the daily weather indexes from January 2007 to April 2007 in Hong Kong. For illustration, the testing data are listed in the Appendix.

For each experimental model, we have conducted ten trials. The average prediction accuracies of the minimum and maximum monthly temperatures by the different models are summarized in Table VII. The comparison shows that the predictive ability of the CIII daily temperature predictor is generally superior to the other three approaches. Although the CIII daily temperature predictor generally needs longer time for tuning the internal regression coefficients by the training data, it can quickly perform the prediction task once the training process is completed. Unlike the other three approaches, the CIII temperature predictor can directly handle the interval data.

TABLE X
HISTORICAL DATA OF THE DAILY WEATHER INDEXES OF MARCH 2007 IN HONG KONG

Day	Temperature range		Humidity range		Rainfall
	Minimum	Maximum	Minimum	Maximum	
1	23.1	25.5	77.0	90.0	0.01
2	15.3	23.7	84.0	91.0	1.3
3	13.7	16.0	53.0	86.0	0.2
4	13.3	17.6	48.0	72.0	0
5	15.5	20.5	55.0	82.0	0
6	17.4	22.0	44.0	72.0	0
7	16.5	19.5	55.0	74.0	0
8	15.1	19.2	61.0	78.0	0
9	15.7	21.0	60.0	86.0	0
10	17.2	23.6	64.0	87.0	0
11	20.5	26.2	55.0	86.0	0
12	21.4	26.1	64.0	89.0	0
13	18.7	21.9	83.0	90.0	0.01
14	18.8	21.1	75.0	88.0	0
15	18.8	21.2	69.0	89.0	0.01
16	18.7	21.9	77.0	89.0	0.01
17	20.2	25.2	80.0	95.0	0.5
18	19.1	23.2	77.0	97.0	0.01
19	17.7	20.5	68.0	78.0	0
20	18.6	22.3	66.0	79.0	0
21	18.5	22.3	65.0	80.0	0
22	18.0	21.2	65.0	84.0	0.01
23	18.1	19.8	71.0	92.0	0.01
24	18.2	19.1	77.0	93.0	0.1
25	17.4	18.2	93.0	98.0	8.3
26	15.9	18.0	84.0	98.0	3.2
27	15.7	17.7	85.0	93.0	0.1
28	16.9	18.8	87.0	96.0	0.1
29	18.0	19.2	92.0	96.0	0.4
30	18.8	21.7	83.0	99.0	89
31	18.3	19.4	88.0	97.0	1.1

TABLE XI
HISTORICAL DATA OF THE DAILY WEATHER INDEXES OF APRIL 2007 IN HONG KONG

Day	Temperature range		Humidity range		Rainfall
	Minimum	Maximum	Minimum	Maximum	
1	18.9	21.7	89.0	97.0	21
2	18.8	21.1	69.0	91.0	0.01
3	16.5	22.0	73.0	90.0	2.2
4	18.9	22.6	79.0	93.0	0.01
5	17.8	20.6	74.0	86.0	0
6	19.4	23.1	78.0	93.0	0
7	21.3	25.0	79.0	94.0	0.2
8	18.0	23.1	68.0	95.0	2
9	18.1	20.7	69.0	86.0	0
10	19.2	23.4	69.0	88.0	0
11	21.2	27.2	68.0	92.0	0
12	22.5	27.0	59.0	93.0	0
13	23.8	28.6	73.0	93.0	0.01
14	20.2	24.8	84.0	97.0	54.3
15	20.0	21.8	83.0	96.0	0.3
16	21.1	23.2	83.0	97.0	2.4
17	21.4	26.8	77.0	97.0	55.4
18	21.4	26.8	45.0	95.0	0
19	22.4	29.2	51.0	89.0	0
20	23.1	27.5	68.0	90.0	0
21	24.1	27.4	76.0	93.0	0
22	25.4	30.7	63.0	88.0	0
23	25.9	29.9	67.0	88.0	0
24	22.3	27.1	79.0	91.0	0.01
25	20.9	23.4	82.0	94.0	0.01
26	22.9	26.8	77.0	93.0	0.01
27	23.3	29.5	61.0	95.0	8.5
28	21.8	24.8	72.0	92.0	0.9
29	22.1	23.6	85.0	94.0	0.01
30	22.9	28.6	75.0	93.0	0

Both the input and output data are allowed to be interval valued. The CIII daily temperature predictor can describe and elicit the interaction among the interval-valued data more precisely than those approaches that preprocess the interval data by separating them into two crisp data (the minimum and maximum values). Another advantage of the CIII daily temperature predictor is that the internal relationships among the predictive attributes (weather indexes) could be explicitly expressed by the values of the signed fuzzy measure μ. It is helpful to meteorologists for figuring out the valuable information for further research and model construction.

VII. CONCLUSION

In this paper, we have established the FCIFI as a generalization of the original Choquet integral. In this fuzzification, both the integrand and the integration result are fuzzy valued. We have conducted the theoretical derivation of the FCIFI based on the extension principle in the fuzzy set theory. Due to this definition, the calculation of the FCIFI has been transformed into that of the CIII. We have given the sound definition of the CIII, which takes an interval-valued integrand and gives an interval-valued integration result as well. We have proposed a calculation scheme with a relevant algorithm to obtain the value of the FCIFI with respect to a fuzzy measure according to the monotonicity of the fuzzy measure. For the FCIFI with respect to a signed fuzzy measure, we have designed a specific GA to search for the interval's terminals of the integration result of the CIII.

We have designed a CIII regression model for the regression problems where interval-valued data are involved. We have derived a double-GA optimization algorithm for computing the internal regression coefficients of the CIII regression model, where one genetic approach is for the parameter optimization and the other approach is for the calculation of the CIII with respect to a signed fuzzy measure. The CIII regression model has been utilized as a daily temperature predictor. In this predictor, the interval-valued weather indexes, such as the temperature range, relative humidity range, and gross rainfall, of several days before the day being considered act as the predictive attributes to forecast the temperature range of the day being considered by a CIII regression model. We have implemented a series of experiments to validate the performance of the daily temperature predictor by real weather records from the Hong Kong Observatory. The results are generally better than those of the existing temperature predictors.

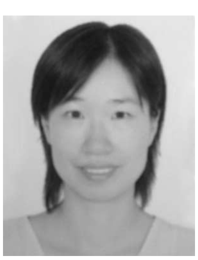
APPENDIX

The testing data in Section VI-B is the collection of the historical data of the daily weather indexes from January 2007 to April 2007 in Hong Kong, which are listed in Tables VIII–XI.

REFERENCES

[1] R. J. Aumann, "Integrals of set-valued functions," *J. Math. Anal. Appl.*, vol. 12, no. 1, pp. 1–12, Aug. 1965.
[2] J. J. Buckley and E. Eslami, *An Introduction to Fuzzy Logic and Fuzzy Sets*. New York: Physica-Verlag, 2002.
[3] S. M. Chen and J. R. Hwang, "Temperature prediction using fuzzy time series," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 30, no. 2, pp. 263–275, Apr. 2000.
[4] G. Choquet, "Theory of capacities," *Ann. Inst. Fourier*, vol. 5, pp. 131–295, 1954.
[5] D. Denneberg, *Non-Additive Measure and Integral*. Boston, MA: Kluwer, 1994.
[6] M. Grabisch and J. M. Nicolas, "Classification by fuzzy integral: Performance and tests," *Fuzzy Sets Syst.*, vol. 65, no. 2/3, pp. 255–271, Aug. 1994.

[7] M. Grabisch, H. T. Nguyen, and E. A. Walker, *Fundamentals of Uncertainty Calculi, With Applications to Fuzzy Inference*. Boston, MA: Kluwer, 1995.
[8] Y.-P. Huang and T.-M. Yu, "The hybrid grey-based models for temperature prediction," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 27, no. 2, pp. 284–292, Apr. 1997.
[9] G. J. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic—Theory and Applications*. Upper Saddle River, NJ: Prentice-Hall, 1995.
[10] K.-S. Leung, M.-L. Wong, W. Lam, Z. Wang, and K. Xu, "Learning nonlinear multiregression networks based on evolutionary computation," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 32, no. 5, pp. 630–644, Oct. 2002.
[11] T. Murofushi, M. Sugeno, and M. Machida, "Non-monotonic fuzzy measures and the Choquet integral," *Fuzzy Sets Syst.*, vol. 64, no. 1, pp. 73–86, May 1994.
[12] E. Pap, *Null-Additive Set Functions*. Boston, MA: Kluwer, 1995.
[13] M. Sugeno, "Theory of fuzzy integral and its applications," Ph.D. dissertation, Tokyo Inst. Technol., Tokyo, Japan, 1974.
[14] Z. Wang, "A new genetic algorithm for nonlinear multiregressions based on generalized Choquet integrals," in *Proc. 12th IEEE Int. Conf. Fuzzy Syst.*, 2003, vol. 2, pp. 819–821.
[15] Z. Wang and G. J. Klir, *Fuzzy Measure Theory*. New York: Plenum, 1992.
[16] Z. Wang, K. S. Leung, M. L. Wong, J. Fang, and K. Xu, "Nonlinear nonnegative multiregressions based on Choquet integrals," *Int. J. Approx. Reason.*, vol. 25, no. 2, pp. 71–87, Oct. 2000.
[17] Z. Wang, R. Yang, P. A. Heng, and K. S. Leung, "Real-valued Choquet integrals with fuzzy-valued integrand," *Fuzzy Sets Syst.*, vol. 157, no. 2, pp. 256–269, Jan. 2006.
[18] K. Xu, Z. Wang, P.-A. Heng, and K.-S. Leung, "Classification by nonlinear integral projections," *IEEE Trans. Fuzzy Syst.*, vol. 11, no. 2, pp. 187–201, Apr. 2003.
[19] Z. Wang, R. Yang, and K. S. Leung, "On the Choquet integral with fuzzy-valued integrand," in *Proc. 11th World Congr. Int. Fuzzy Syst. Assoc.*, 2005, pp. 433–437.
[20] R. Yang, Z. Wang, P. A. Heng, and K. S. Leung, "The calculation of Choquet integrals with fuzzy-valued integrand," in *Proc. Int. Conf. Inf. Proc. Manage. Uncertainty Knowledge-Based Syst.*, 2004, pp. 1263–1270.
[21] R. Yang, Z. Wang, P. A. Heng, and K. S. Leung, "Fuzzy numbers and fuzzification of the Choquet integral," *Fuzzy Sets Syst.*, vol. 153, no. 1, pp. 96–113, Jul. 2005.
[22] R. Yang, Z. Wang, P. A. Heng, and K. S. Leung, "Classification of heterogeneous fuzzy data by Choquet integral with fuzzy-valued integrand," *IEEE Trans. Fuzzy Syst.*, vol. 15, no. 5, pp. 931–942, Oct. 2007.
[23] L. A. Zadeh, "Soft computing and fuzzy logic," *IEEE Softw.*, vol. 11, no. 6, pp. 48–56, Nov. 1994.



Rong Yang received the B.Sc. (Eng.) degree in electrical engineering from Southeast University, Nanjing, China, in 1998, the M.Phil. degree in electronic and electrical engineering from The Hong Kong University of Science and Technology, Kowloon, in 2000, and the Ph.D. degree in computer science from The Chinese University of Hong Kong, Shatin, in 2005.

She is currently an Assistant Professor with the Department of Automatic Science, College of Mechatronics and Control Engineering, Shenzhen University, Shenzhen, China. Her research interests include fuzzy theory, nonlinear integrals, nonlinear optimization, pattern recognition, soft computing techniques, and data mining.

Zhenyuan Wang received the degree from Fudan University, Shanghai, China, and the Ph.D. degree from the State University of New York at Binghamton.

He was a Visiting Professor with the State University of New York at Binghamton, New Mexico State University, Las Cruces, and the University of Texas at El Paso and a Research Fellow with The Chinese University of Hong Kong, Shatin. He is currently a Professor with the Department of Mathematics, University of Nebraska, Omaha. He is a member of the Editorial Board of *Fuzzy Sets and Systems* and the *Journal of Fuzzy Mathematics*. He currently serves as an Associate Editor of the *Journal of Intelligent and Fuzzy Systems*. He is the author or coauthor of more than 100 papers published in international journals and conference proceedings and is the coauthor of one book. His research interests include fuzzy measure theory, nonlinear integrals, nonlinear optimization, soft computing techniques, and data mining.

Pheng-Ann Heng (S'90–M'92) received the B.Sc. degree from the National University of Singapore, Singapore, in 1985 and the M.Sc. degree in computer science, the M.A. degree in applied mathematics, and the Ph.D. degree in computer science from Indiana University, Bloomington, in 1987, 1988, and 1992, respectively.

He is currently a Professor with the Department of Computer Science and Engineering, The Chinese University of Hong Kong (CUHK), Shatin. In 1999, he set up the Virtual Reality, Visualization and Imaging Research Centre, CUHK, and serves as the Director of the Centre. He is also the Director of the CUHK Strategic Research Area in Computer Assisted Medicine, which was jointly established by the Faculty of Engineering and the Faculty of Medicine in 2000. His research interests include virtual reality applications in medicine, visualization, 3-D medical imaging, user interface, rendering and modeling, and interactive graphics and animation.

Kwong-Sak Leung (M'77–SM'89) received the B.Sc. (Eng.) and Ph.D. degrees from the Queen Mary College, University of London, London, U.K., in 1977 and 1980, respectively.

He worked as a Senior Engineer on contract research and development with ERA Technology and later joined the Central Electricity Generating Board to work on nuclear power station simulators in the U.K. In 1985, he joined the Department of Computer Science and Engineering, The Chinese University of Hong Kong, Shatin, where he is currently a Professor of computer science and engineering. He is a member of the Editorial Board of *Fuzzy Sets and Systems* and serves as an Associate Editor of the *International Journal of Intelligent Automation and Soft Computing*. He is the author or coauthor of more than 200 papers published in international journals and conference proceedings and of two books in fuzzy logic and evolutionary computation. His research interests are in soft computing, including evolutionary computation, parallel computation, probabilistic search, information fusion and data mining, fuzzy data, and knowledge engineering.

Prof. Leung is a member of the Institution of Electrical Engineers and the Association for Computing Machinery, a Fellow of the Hong Kong Institution of Engineers, a Distinguished Fellow of the Hong Kong Computer Society, and a Chartered Engineer. He has been the Chair and a member of many program and organizing committees of international conferences.